Expression Robust 3D Face Landmarking Using Thresholded Surface Normals

Jiangning Gao[a,*], Adrian N. Evans[b]

[a] *Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden. Email: jiangning.gao@imbim.uu.se*
[b] *Department of Electronic and Electrical Engineering, University of Bath, Bath, United Kingdom. Email: A.N.Evans@bath.ac.uk*

Abstract

3D face recognition is an increasing popular modality for biometric authentication, for example in the iPhoneX. Landmarking plays a significant role in region based face recognition algorithms. The accuracy and consistency of the landmarking will directly determine the effectiveness of feature extraction and hence the overall recognition performance. While surface normals have been shown to provide high performing features for face recognition, their use in landmarking has not been widely explored. To this end, a new 3D facial landmarking algorithm based on thresholded surface normals maps is proposed, which is applicable to widely used 3D face databases. The benefits of employing surface normals are demonstrated for both facial roll and yaw rotation calibration and nasal landmarks localisation. Results on the Bosphorus, FRGC and BU-3DFE databases show that the detected landmarks possess high within-class consistency and accuracy under different expressions. For several key landmarks the performance achieved surpasses that of state-of-the-art techniques and is also training free and computationally efficient. The use of surface normals therefore provides a useful representation of the 3D surface and the proposed landmarking algorithm provides an effective approach to localising the key nasal landmarks.

Keywords: 3D Face Landmarking, Surface Normals

*Corresponding author

1. Introduction

The 3D face surface is very distinctive over different identities and is widely used in biometric identification. Using 3D facial data to understand the characteristics of the human face in the 3D domain has been shown to have much promise for improving the overall recognition performance [1][2][3]. Landmarking is an important stage of many face recognition algorithms [4] and, as such, the accuracy of the landmarking directly affects the effectiveness of subsequent feature extraction and matching. Inconsistent landmarking may adversely affect the within-class similarity and between-class dissimilarity of facial features. In addition, for real-time automatic face recognition systems, the landmarking method should also be computationally efficient. Many 3D face landmarking algorithms have been proposed for region-based feature extraction, alignment or other 3D face analysis. Current landmarking algorithms are often based on the depth information calculated from the original 3D captures, which possess fine accuracy and within-class consistency but might not be accurate enough for the other 3D surface represenations, for example surface normals (SN).

In the recent years, many recognition algorithms have been proposed to extract discriminative features from the SN on the 3D face surface, which have great potential to produce better recognition performance than using the depth information [2][5][6]. In particular, using region-based feature extraction algorithms on the SN has been widely explored, in order to overcome the challenges caused by expression variations [2][5]. However, most existing landmarking approaches are based of the depth information, which can increase the variance in the SN by inaccurate landmarks. For example, two closely located landmarks on the facial surface will typically have similar depth values but may have very different surface normals. Therefore, an accurate and consistent landmarking approach is essential for region-based feature extraction using SN. To extract the discriminative features from the SN a promising approach is to use the information in the SN to refine and improve existing landmarking approaches.

To this end, a new landmarking strategy that is based on utilizing the SN

is proposed. As an alternative to using depth maps, SN have the potential to provide a comprehensive description of local facial details and geometrical information. This paper demonstrates that many of the local structures required for landmarking are evident in the SN and can be easily be localised by simple thresholding of the SN maps. In the approach, the nasal bridge and candidate alar grooves are first robustly detected by thresholding the SN components (SN_x, SN_y and SN_z). These detected components are then used to refine the location of the nose tip, previously found in the depth map, and correct the rotational alignment in the yaw direction. Also, using the nasal bridge, a coarse to fine roll rotation calibration is applied to further reduce the pose variations; this step is found to be beneficial even though the use of SN is relatively pose invariant. Finally, seven landmarks on the nasal region are robustly detected on the well-aligned 3D captures by utilising the thresholded SN and facial profile signatures.

In addition, to improving the overall consistency and accuracy of the detected landmarks, the proposed approach can address pose variations in the yaw and roll directions and inaccurate nose tip localisation. The Bosphorus [7], FRGC [8] and BU-3DFE [9] databases contain many high resolution 3D captures and different expressions and are used for the performance evaluation, which shows the robustness and high performance of the proposed landmarking algorithm.

2. Related Work

A comprehensive summary of widely used landmarking algorithms in both academia and industry is provided by Creusot et al. [10] in which three categories of landmarking algorithms are defined: curvature based descriptors, facial profile signatures and machine learning based methods. Using the curvature descriptor maps to find the salient points on the 3D captures is widely used by many researchers [11][12][13][14][15][16][17]. For example, HK classification thresholds the mean (H) and Gaussian (K) curvatures to find the most salient

landmark candidates and is one of the most popular landmarking methods. In some approaches, after the HK classification the landmark are defined as the extrema points in given directions [11][13]. However, the majority of algorithms utilising HK classification are sensitive to yaw, pitch and roll rotations, even though the H and K curvatures of a surface are rotation invariant [16]. In [17], a cropped face template containing three main landmarks (nose tip and eyes) is used for validation. This approach also suffers from variations under yaw and pitch rotations.

Instead of employing HK classification, the shape index (SI) and curvedness [18], computed from principle curvatures, can provide effective maps for landmarking. A popular approach is to detect the nose tip using the largest SI and, similarly, the two inner eye corners can be found by detecting the smallest SI values[12]. In most approaches, the nose tip is always located at the position of maximum depth on the largest convex region. Lu and Jain [19] detected seven corresponding points on 3D faces using shape index from range images and cornerness from intensity images.

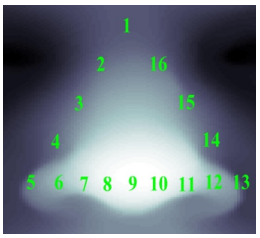
Another popular approach is to find facial profiles and transections of the 3D point cloud and use these to localise the nose tip, root and corners. For example, Faltemier et al. proposed the use of rotated profile signatures to extract the contours of a mesh, which is then moved through varying at angles until it matches a previously learned nose profile signature [20]. Combined with features extracted from the mean and Gaussian curvatures, Segundo et al. [15] have also used transverse slices to detect the nose tip and the nose corners in 3D faces of the FRGC and BU-3DFE databases.

Instead of using a heuristic method for landmark detection, machine learning based approaches can be employed [10][21][22]. Creusot et al. proposed a machine-learning based approach to detect 14 corresponding landmarks on 3D faces. to build a strong feature vector from a 3D mesh by combining different local surface descriptors [10]. To best separate each landmark from its surrounding region within the set of training data, different local surface descriptors and optimal function (LDA and AdaBoost) are used. Perakis et al. [23] proposed

to detect the keypoints by using the shape descriptors and then the candidate landmarks are identified and labelled by matching them with a Facial Landmark Model of facial anatomical landmarks. Combined with both 2D (edge response) and 3D (SI and spin image) facial feature descriptors, a generalized face landmark detection was developed in [24].

Using a sparse set of correspondences, some methods create morphable models and then extend them to dense correspondences. Landmark localization is performed on the segmented face by finding the transformation that minimizes the deviation of the Point Distribution Model from the mean shape. Sukno et al. proposed an automatic landmarking algorithm that integrates nonrigid deformation with the ability to handle missing points [25]. This approach first generates sets of candidate locations from feature detectors and then performs combinatorial search, constrained by a flexible shape model. More recently, Gilani et al. proposed a method to automatically establish dense correspondences between a large number of 3D faces [26]. The existing correspondences are triangulated and then expanded iteratively by matching points of distinctive surface curvature along the triangle's edges. An algorithm is proposed for morphing the constructed deformable model to fit unseen faces. Morphable model approaches are able to generate consistent, dense fields of landmarks but typically require extensive training and are computationally expensive.

Emambakhsh et al. proposed an effective curvature and depth based landmarking algorithm for the nose tip, root and two alar grooves [27], from which the set of 16 nasal landmarks shown in Figure 1a were defined. In this approach, the convex regions were found by thresholding the SI [27][28][29] and the nose tip was located at the centroid of the largest convex region of the face. The nasal root was found by identifying the maximum of the minima of a set of nasal curves passing through the nose tip. However, this method is not always robust for some scenarios. For example, the shape of human nose is diverse and the nasal root is not necessary located at the minimum of each curve, which results in failed detection for some subjects. The location of nasal alar grooves is also sensitive to residual noise.



(a) 16 nasal landmarks [27]

(b) Binary map of nasal region

Figure 1: 16 nasal landmarks and the binary map for the key landmarks detection. (a) shows 16 landmarks found in [27], where the nose tip, root (upper) and two alar grooves (lower left and right) are the key landmarks. (b) binary map of the nasal region after surface normals thresholding indicates the candidate regions of all the key landmarks.

In practice, most curvature based methods are sensitive to residual noise, pose variations or occlusions by glasses. In addition, denoising and accurate pose alignment are also required. One simple and effective way to address those problems is to add differential geometry constraints to optimize the locations of the candidate landmarks. As a first order surface differential quantity, the SN define the orientation of a surface and provide an effective way to localise landmarks. An example of surface normals thresholding is shown in Figure 1b, which illustrates the candidate regions of the nasal root and the alar grooves.

3. Landmarking Using Surface Normals

An overview of the proposed landmarking algorithm is presented in the flowchart in Figure 2. In the initial steps, following [27] the nose tip is first detected by thresholding the SI map and then the initial nasal region is cropped using the nose tip. After Principal Component Analysis (PCA) based alignment, most pose variations are corrected and the tip is further localised. However, some rotations still remain, especially in the yaw and roll directions, resulting from unexpected facial occlusions or special facial structures, which are hard to correct using PCA based alignment. Therefore, in this section a novel landmarking algorithm is proposed, which aims to address the remaining pose variations

using the nasal geometrical structure and the detected nasal landmarks, as well as further improve the accuracy and consistency of nasal landmarks.

Figure 2 shows the stages of the proposed algorithm and a brief description of its main steps is given below. Using the thresholded SN maps, the nasal bridge is first localised and used to provide a coarse roll correction. A new nose tip detection method is then applied, which uses the nasal bridge and nose profile signature. Using the refined tip location, candidate alar grooves regions are found by thresholding the SN and then used to correct self-yaw rotations, relocate the tip and provide a reference map for the final alar grooves localisations on the depth map. The nasal root and subnasal are localised in a similar manner to that used for the tip detection. As the nasal root detection is sensitive to roll rotation, a fine roll rotation pose correction is proposed that estimates the rotation angle using the detected root and tip. Finally, seven nasal landmarks (the tip, root, subnasal, alar grooves and eye corners) are localised on the well aligned 3D nose. Each of these stages are fully described in the following subsections.

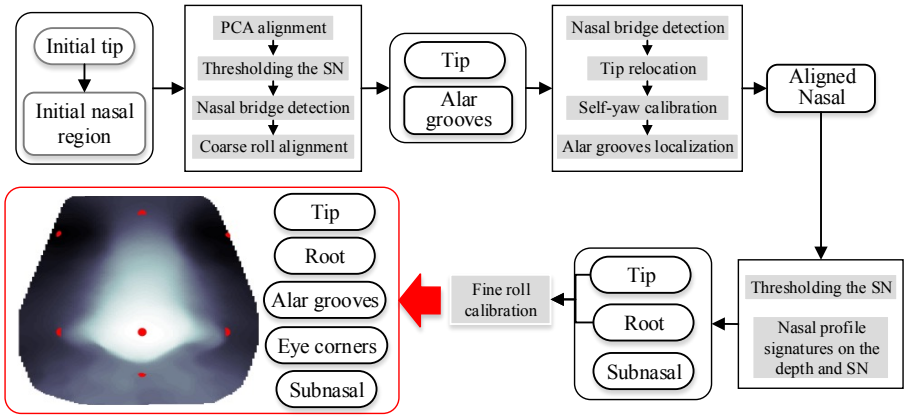


Figure 2: The flow chart of the proposed landmarking algorithm: Using the direct outputs of PCA based alignment, the cropped nasal region and the initial tip, the thresholded SNs and nasal bridge are first found for the coarse roll and self-yaw calibration. Then, the location of the root and tip are used for further roll calibration. Finally, the nose tip, root, two alar grooves and subnasal are localised on the aligned nasal region. Eye corners can be localised by thresholding the SNs.

3.1. Nose Tip Detection and Coarse Roll Calibration

As shown in the example capture in Figure 3a, some pose variations still remain after the PCA based alignment. These are mainly in the roll direction and are predominantly caused by occlusions due to hair, which cannot be addressed by PCA based alignment. The first step in this stage is a coarse roll rotation calibration, to approximately align the nasal region before landmarking. As one of the most salient parts on the nasal region, the nasal bridge is found and used to estimate the coarse roll rotation angle.

3.1.1. Nasal Bridge Detection

The nasal bridge is a distinct geometrical structure on the facial surface and as its centre should be vertical in the frontal view face its location can indicate the degree of roll variations. Therefore, using this prominent structure can provide a good frame of reference to rapidly correct approximate facial roll variations. Compared to using the depth information, thresholding the SN_x map provides an effective way to detect the nasal bridge as it has SN_x values that are close to zero. The nasal bridge region can be easily detected by thresholding the SN_x map to give TSN_x. The TSN_x region is a matrix with the same size as SN_x and is found by,

$$TSN_x = \{TSN_{x_{i,j}} | TSN_{x_{i,j}} = f_x(i,j)\} \tag{1}$$

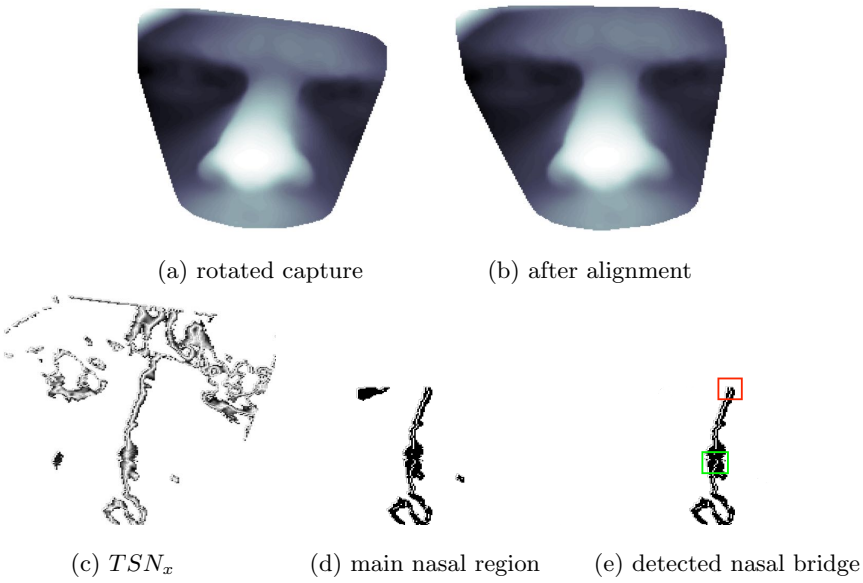
where

$$f_x(i,j) = \begin{cases} 1 & if\ (SN_x(i,j) < T_x) \\ 0 & otherwise \end{cases} \tag{2}$$

An example TSN_x region for a capture from the Bosphorus database is shown in Figure 3c.

To detect the nasal bridge in thresholded maps such as that shown in Figure 3c, the main nasal region is first cropped using the location of tip as shown in Figure 3d. The nasal bridge is located as the largest connected component of the main nasal region, see Figure 3e. For some 3D captures, the extracted main nasal region is not fully connected, mainly due to noisy points or varying minor

structures for some identities. Therefore, a morphological closing is applied to the thresholded SN_x before detecting the largest connected component.



(a) rotated capture (b) after alignment

(c) TSN_x (d) main nasal region (e) detected nasal bridge

Figure 3: Coarse roll rotation calibration using the detected nasal bridge: (a) An example misaligned capture from the Bosphorus database. (b) Results of calibration by the detected nasal bridge region. The candidate nasal bridge is given by the TSN_x in (c) and the main nasal region shown in (d) is extracted on the basis of the location of nose tip. The largest connected area shown in (e) is then detected from (d). Using the estimated top (red rectangle) and bottom (green rectangle), the rotation angle can be estimated.

3.1.2. Coarse Roll Calibration

As can be seen from Figure 3e, the detected nasal bridge is an irregular structure whose vertical deviation is hard to estimate. In this section, an efficient method to rapidly estimate the rotation angle of the y axis (α_1) is proposed, using the centroids of the top (top_x, top_y) and bottom ($bottom_x$, $bottom_y$) of the nasal bridge shown in red and green in Figure 3e.

The previously detected nose tip is used to help localise the two centroids. On the detected nasal bridge map shown in Figure 3e, the position of the nose tip detected in the alignment phase is used to indicate the horizontal range

of the top and the bottom can be localised by the neighbouring points of the tip. The top of the detected nasal bridge, shown as the red rectangle, is used to estimate the top. Using the nasal bridge, all the pose variations in the roll direction can be detected and corrected. As shown in Figure 3b, the rotated capture $F_1(x_1, y_1, z_1)$ is computed from the original capture $F(x, y, z)$ with the rotation angle α_1 in the x-y plane, given by,

$$\alpha_1 = tan^{-1}\frac{top_x - bottom_x}{top_y - bottom_y} \tag{3}$$

The new aligned coordinates are then given by

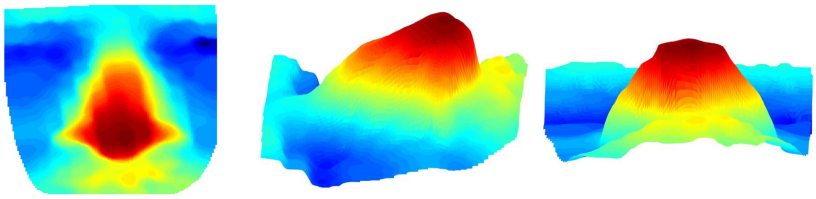
$$\begin{cases} x_1 = x \times cos\alpha_1 - y \times sin\alpha_1 \\ y_1 = x \times sin\alpha_1 + y \times cos\alpha_1 \\ z_1 = z \end{cases} \tag{4}$$

3.1.3. Nose Tip Detection

Curvature and SI are widely used for the nose tip detection as they can effectively detect the convex regions, with the tip located in the largest convex region. However, the main challenge of this approach is how to accurately locate the tip within this region. In many landmarking algorithms, finding the centroid or the highest point are popular approaches and can produce accurate and consistent tip locations. However, these approaches are not always reliable and can be adversely affected when unexpected problems occur during the data acquisition or the target captures contain particular nasal structures, which persist after application of the denoising algorithms. Figure 4 gives an example of this kind of capture, where the centroid of the convex region gives the incorrect tip position. In these cases, simply considering the depth information of nasal region might not necessarily produce an accurate tip position.

Standard denoising approaches will not address these problems as they can cause the fine details to be lost in other, unaffected captures. Also, identifying affected captures is a challenging issue. Instead of increasing the aggressiveness of the denoising using the TSN_x map, for example that in Figure 3c, the nasal

profile shown in Figure 5 can provide a novel and effective approach for nose tip detection. First, for a frontal view facial capture, the tip position in the x axis, Tip_x, is determined using the nasal bridge, which can be estimated by averaging its neighbouring points on the nasal bridge within the green box shown in Figure 3e. Using this estimate the nasal profile is found by intersecting a plane that is perpendicular to the x-y plane with the nasal surface. This results in a curve such as that shown in Figure 5. This curve is further smoothed to remove noise and the tip position in the y axis, Tip_y, is located at the highest point of the curve.



(a) Frontal view (b) Side view (c) Upward view

Figure 4: An example facial capture from the Bosphorus database (shown in three views) that suffers unexpected problems during the data acquisition, and fails in the tip detection when only the depth information is used. In Bosophorus database, a few subjects contain this kind of captures, which mainly results from some unexpected problems occurred during data acquisition.

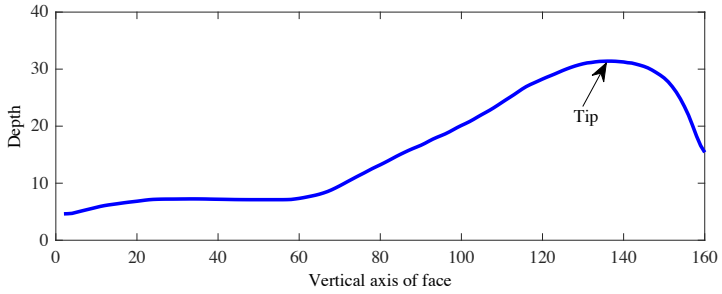


Figure 5: Tip localisation by nasal bridge and profile. Using Tip_x found on the nasal bridge, a curve on the nasal surface is helped to localise the final tip. After smoothing the noisy points on the profile curve, Tip_y is found at the highest point.

3.2. Alar Grooves Detection and Self-Yaw Rotation Calibration

In most landmarking methods employing the depth map, the accuracy of the alar grooves' locations can be adversely affected by noisy points or an inaccurate nose tip position. In this section, thresholded SN(TSNs) are used to create an accurate candidate alar grooves map, to further refine the tip location and to correct yaw rotations. The alar grooves are finally localised on the depth map. The main steps of this process are shown in Figure 6 and are described below.

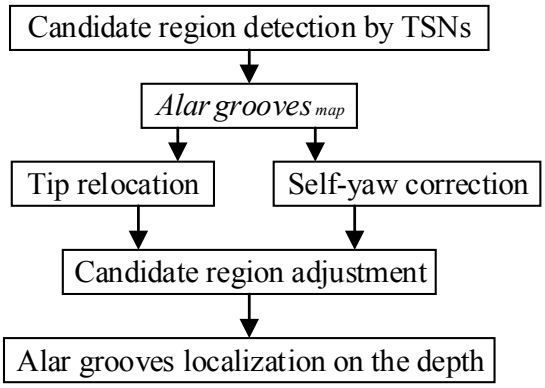


Figure 6: The flow chart of the proposed alar grooves detection approach. The candidate alar grooves regions are first detected Section 3.2.1. Using the candidate regions, the nose tip is further refined Section 3.2.2 and the self-yaw rotation is corrected Section 3.2.3. After the candidate regions adjustment Section 3.2.4, the alar grooves are finally localised on the depth map Section 3.2.5.

3.2.1. Candidate Region Detection by TSNs

The SN of the nasal alar grooves and nostrils have higher SN_x and SN_y values and very small values for SN_z when compared to the SN of other nasal and adjoining cheek regions. Therefore, the outline of the nasal alar grooves and nostrils is very salient in comparison with other regions around the nasal region and the candidate alar grooves can be obtained directly from the SN_z map. By thresholding the SN_z map, the initial candidate region is found, see Figure 7. This approach is more straightforward than using a combination of the SN_x and SN_z maps. Although some small patches on the upper nasal region are

also present in the thresholded map, these can be excluded by using the nose tip as a reference. The $Alar\ grooves_{map}$ is a matrix with the same size as SN_z and is given by

$$Alar\ grooves_{map} = \{Ag_{i,j} | Ag_{i,j} = f_z(i,j)\} \tag{5}$$

where, given SN_z,

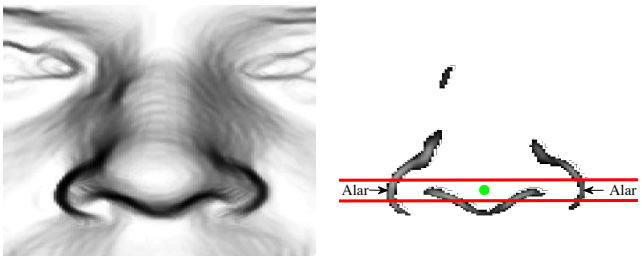
$$f_z(i,j) = \begin{cases} 1 & if\ (SN_z(i,j) < T_z) \\ 0 & otherwise \end{cases} \tag{6}$$

For the example capture in Figure 7, the location of the nose tip [green point in Figure 7b] detected in the previous stage is used to define two red lines at a fixed vertical separation from the original nose tip, which is detected from the depth map.

The resulting horizontal region between the two red lines of height i pixels gives the target region in which to find the candidate alar grooves. For each line in this region, the initial candidate alar grooves points, $init\_Left_i$ and $init\_Right_i$, are obtained by finding the left and right extrema, respectively. However, self-occlusion caused by pose variations and inaccurate nose tip location can greatly reduce the landmarking accuracy. To address this problem, nose tip relocation, self-yaw calibration and interpolation of missing points are proposed.

3.2.2. Nose Tip Relocation

In the last section, the position of nose tip (J_{tip}) is localised using the nasal bridge to provide a coarse and reliable estimate. To further refine the detected nose tip position, and calibrate the remaining yaw rotations, the candidate alar grooves are used to define a new nose tip location (J_new_{tip}). Combined with these two types of estimations of the nose tip, the rotation angle in the yaw direction is estimated from their differences. The initial candidate points, $init\_Left_i$ and $init\_Right_i$, can be presented as i pairs of candidate points. Most pairs on the frontal view captures are relatively symmetrically distributed



(a) (b)

Figure 7: Candidate alar grooves detection: (a) Example SN_z map of nasal region, where darker shades denote lower values. The alar grooves and nostrils are more salient than the other parts on the SN_z map as their values are higher than the other parts. (b) The thresholded SN_z map used to find the candidate alar groove region. Based on the location of nose tip (green point), two horizontal red lines are used to determine the upper and lower boundary of the target region.

on each side of the nose. Therefore, it is possible to further refine the location of the nose tip using the detected pairs. The set of pairs can be found by

$$\begin{cases} Dist\_Left_i = J_{tip} - init\_Left_i \\ Dist\_Right_i = init\_Right_i - J_{tip} \\ D_i = |Dist\_Left_i - Dist\_Right_i| \\ D_s = min(D_i(Dist\_Left_i \neq 0 \wedge Dist\_Right_i \neq 0)) \end{cases} \tag{7}$$

where the distances between the initial candidate points ($init\_Left_i$ and $init\_Right_i$) and current nose tip (J_{tip}), $Dist\_Left_i$ and $Dist\_Right_i$, are first calculated. The absolute distance, D_i, is then obtained by calculating the differences between $Dist\_Left_i$ and $Dist\_Right_i$.

Some initial candidate pairs with an invalid distance are first removed and the pairs with minimum distance difference, D_s, are used to predict the new nose tip. Also, the nose tip is not adjusted when the value of D_s is smaller than the threshold T_s, which means that the current nose tip only possesses a small offset when compared to the estimation of the pairs of points. These situations often arise because the human nose is not strictly symmetric and it is therefore not necessary to address them. However, if the value of D_s is larger than T_s,

which means the offset of current nose tip is large, the current tip should be relocated by calculating the average distance of the existing set of pairs with low values of D_i, using

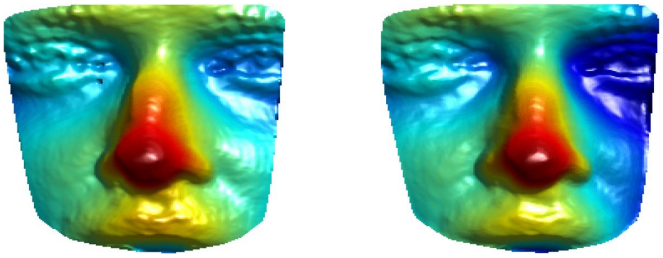
$$J\_new_{tip} = \begin{cases} \lceil \frac{1}{2n} \times \sum_{i=1}^{n} init\_Left_i + init\_Right_i \rceil & D_s > T_s \\ J_{tip} & D_s \leq T_s \end{cases} \tag{8}$$

Therefore, using the candidate alar grooves points, nose tip locations with large deviations can be approximately calibrated and further used for yaw rotation calibration. Noisy data remaining on the facial surface is typically the main reason for these less accurate tip locations. However, some other captures fail in the tip detection due to their intrinsic nose characteristics, for example self-yaw rotated noses. This will be further described and addressed in the next section

3.2.3. Self-yaw Rotation Calibration

For the majority of frontal view facial captures, the position of nasal bridge is in the centre of nasal region and should also contain the nose tip. Therefore, if the estimated J_new_{tip} is not within the nasal bridge region, the captures must contain the yaw rotation. Actually, in the PCA based alignment, the majority of pitch and yaw rotations have been successfully addressed. However, the human nose is not strictly symmetric and some identities have larger pose variations of the nasal region in the yaw direction. An example nose with self-yaw rotation is shown in Figure 8a. The main drawback of such structure is that it can cause the x axis to be reversed after alignment, which greatly decreases the within-class consistency of the landmarking.

This kind of yaw rotation presents more difficulties to the alar grooves localisation: the rotated nose will cause the self-occlusion on one side of the nose corner and it is also hard to detect local features on the SN maps. In general, in most frontal views, noses possess a relatively symmetric structure centred on the nose bridge, which helps the design of an effective landmarking strategy based on symmetry. Therefore, individuals with self-yaw variations can increase



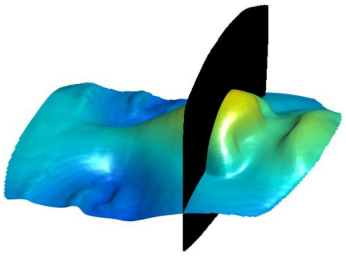
(a) Self-yaw rotation (b) Corrected nasal region

Figure 8: An example capture with self-yaw rotation from the Bosphorus database and its calibrated result. The colour changes shown on the nasal region of (a) demonstrate that the nose suffers the yaw rotation, which is not caused by the facial pose variations as the cheeks regions are shown as frontal.

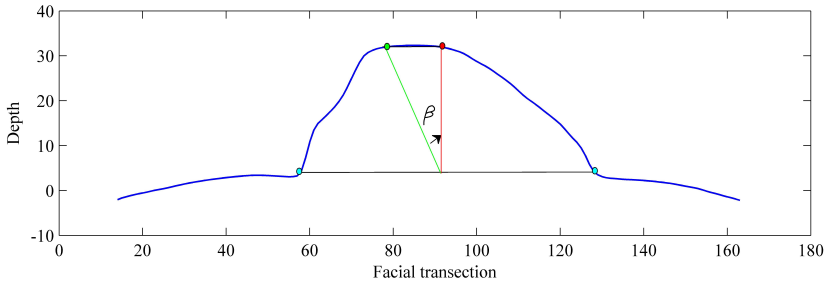
the computational complexity of the alar grooves detection.

To address the self-yaw rotation, the location of the previously detected nose tip, J_{tip} and J_new_{tip}, is initially used to detect the rotation. Then, the detected nasal bridge, such as the example shown in Figure 3c, and candidate alar grooves maps can help address the self-yaw variations. The nasal bridge is one of the most salient geometrical structures on the face surface and its vertical orientation is sensitive to facial rotations in the yaw direction. The detection of the nasal bridge is complicated using depth information but the nasal bridge is more salient on the SN_x map and can be robustly detected by thresholding the SN_x map (TSN_x), even when the whole nasal region has self-yaw rotations.

When combined with the candidate alar grooves ($init\_Left$ and $init\_Right$) matrix obtained in the previous section, which provide a reliable prediction of the nose corners, the yaw rotation angle (β) is estimated by n valid pairs of corner candidate points and their corresponding nasal bridge points ($Nasal_{bridge}$) shown as green in Figure 9b. The rotated capture $F_2(x_2,\ y_2,\ z_2)$ is computed from $F_1(x_1,\ y_1,\ z_1)$ by applying a rotation angle correction β in the x-y plane.



(a) Facial transect



(b) Self-yaw rotation estimation

Figure 9: An example of facial transect passing through the nose tip: Using planes that are perpendicular to the x-y plane results in a series of intersection curves shown in blue on the nasal surface. The rotation angle can be estimated by the detected points on this curve in terms of the tip (red) that is computed by the candidate alar grooves (cyan) and bridge point (green).

The rotated coordinates (x_2, y_2, z_2) are given by

$$\begin{cases} x_2 = x_1 \times cos\beta - z_1 \times sin\beta \\ y_2 = y_1 \\ z_2 = x_1 \times sin\beta + z_1 \times cos\beta \end{cases} \tag{9}$$

To calculate the rotation angle, n facial transects around the nasal tip are found by intersecting planes that are perpendicular to the x-y plane with the face surface. Figure 9 illustrates an example of the transect passing through the nose tip, resulting in an intersection on the 3D facial surface shown in blue. On this transect, there are two nose corners (marked in cyan in Figure 9) which are the candidate alar grooves, a bridge point (green) and the centre point (red).

The angle β is obtained by calculating the rotation angle from the bridge point to the tip by

$$\beta = \frac{1}{n} \times \sum_{i=1}^{n} tan^{-1} \frac{Nasal_{bridge}(i) - \frac{1}{2} \times (init\_Right(i) + init\_Left(i))}{Depth(Nasal_{bridge}(i))} \quad (10)$$

After the coarse roll and self-yaw calibrations, the nose tip is relocated using both the neighbouring points found on the nasal bridge and the prediction by the candidate alar grooves points.

3.2.4. Reconstruction of Missing Points and Candidate Region Adjustment

Noise or self-occlusion can result in some missing or incorrectly detected points in the candidate alar grooves region shown in Figure 9b. To address these problems, using the newly adjusted nose tip (J_new_{tip}), the distances from the candidate points to the tip are recalculated and the missing points are reconstructed from the points from the same lines or interpolated by the neighbouring points, on the basis of the threshold T_c. According to the pairs previously used to predict the new nose tip, T_c is found by calculating the average distance from the candidate points to the new nose tip using,

$$\begin{cases} Dist\_Left_i = J\_new_{tip} - init\_Left_i \\ Dist\_Right_i = init\_Right_i - J\_new_{tip} \\ D_i = |Dist\_Left_i - Dist\_Right_i| \\ D_s = min(D_i(Dist\_Left_i \neq 0 \wedge Dist\_Right_i \neq 0)) \\ T_c = \lceil \frac{1}{m} \times \sum_{i=1}^{m} Dist\_Left_s(i) \rceil \end{cases} \quad (11)$$

where $Dist\_Left_i$ and $Dist\_Right_i$ are the distances between the initial candidate points and current nose tip. The absolute distance, D_i, is obtained by measuring the difference between $Dist\_Left_i$ and $Dist\_Right_i$, which helps find the most appropriate and symmetric pairs, (D_s), where s indicates the selected pairs for the missing points reconstruction.

For each missing or incorrectly detected point, the distance between its symmetrical point in the same line and the nose tip is first compared with T_c to

determine whether it is reconstructed by the pair in the same line or from its neighbouring points. The correction methods of different kinds of missing or false detected points are shown in Figure 10.

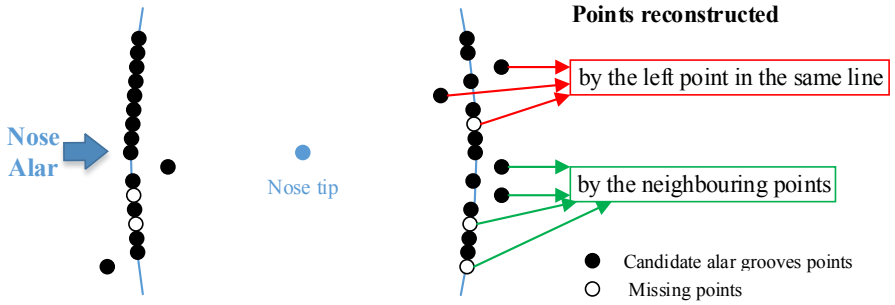


Figure 10: Missing alar points reconstruction and candidate alar grooves adjustment. In this example, each missing or false detected point on the right alar, $Dist\_Right_i$, is reconstructed on the basis of the pair on the same line or the neighbouring points. If the pair on the left is well detected, which is measured by T_c, the point on the right can be reconstructed directly by this pair and the location of nose tip. Otherwise, the point on the right is interpolated from its neighbouring points.

3.2.5. Alar Grooves Localization on the Depth Map

By using the thresholded SNs maps and the location of the nasal bridge, the resulting alar grooves provide an accurate reference map for determining the final alar grooves landmarks on the depth map. The depth information is further applied to produce more accurate alar grooves on the 3D facial surface. More specifically, the candidate alar grooves points shown in cyan in Figure 9b are first localised on the facial transection to give an approximate location on the surface normal maps.

To refine the alar grooves more locally on the depth map, the facial surface curve resulting from intersecting a plane, perpendicular to the x-y plane, is used. In Figure 9b, the two alar grooves should be located at the inflexion points on the facial curves. In a similar manner to the nose tip localisation shown in Figure 5, the most effective way to detect these points is using the rotated curves. Specifically, from the nose tip, the facial curve is first divided

into two halves and the left and right halves are then rotated clockwise and counter-clockwise in the x-y plane, respectively. The two alar grooves are then located at the local minima of two halves.

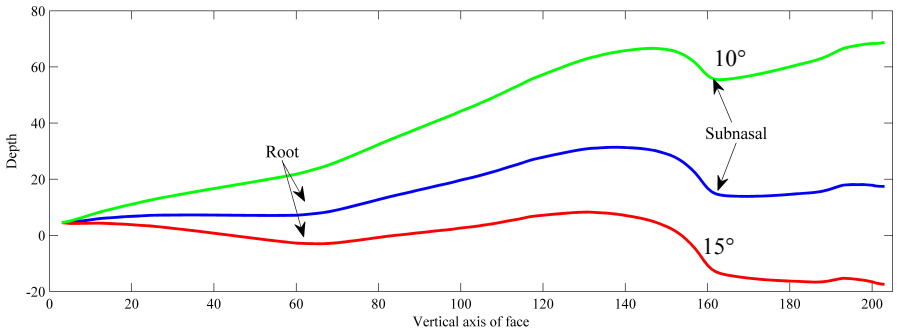
3.3. Fine Roll Rotation Calibration and Root Localization

3.3.1. Nasal Root and Subnasal Localization

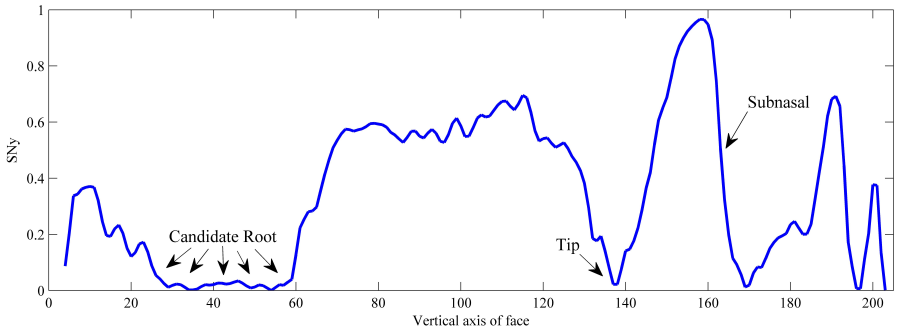
Nasal root localisation is the most challenging step in the landmarking process. It is hard to define an accurate root location as the root area is relatively flat for many human noses. Using a reference area provided by thresholded surface normal maps can successfully estimate the candidate nasal root position. However, simply calculating the centre of this area is not guaranteed to produce an accurate root position. Therefore, using a similar approach to the nose tip adjustment, the position of nasal root and subnasal in x axis are first narrowed by using the detected nasal bridge (for example see Figure 3d) as these two landmarks should be located within the nasal bridge. The fusion of both the thresholded SN maps and the nasal bridge produces a relatively accurate horizontal positions of the root and subnasal, respectively.

As shown in Figure 11, facial profile curves generated on the depth and SN_y maps can be used to find the positions of the nasal root and subnasal in the y axis. These two landmarks should be located at the local minima of the facial profile curves on the depth map. However, it is hard to find the local minima on the profile curves of some captures, for example of the blue curve shown in Figure 11a. To address this problem, the clockwise rotated curve with the rotation angle 10° shown in red in Figure 11a has been proposed to find the local minima and finally used to localise the root [2]. In a similar manner, the subnasal is located at the local minima of the counter clockwise rotated green curve with rotation angle 15°, as shown in Figure 11a.

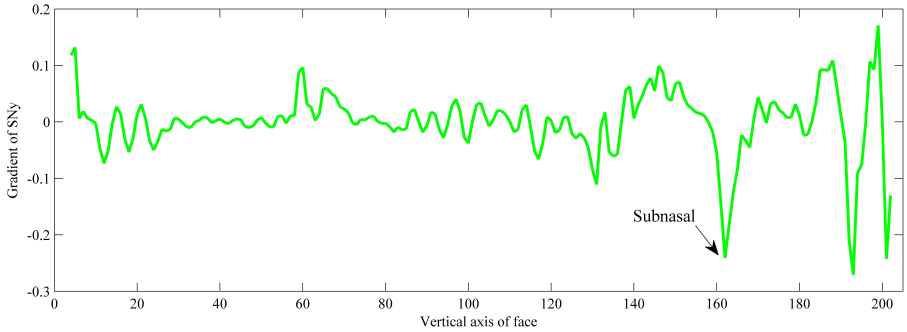
The facial profiles generated on the SN maps also help localise the nasal root and subnasal. The nasal root is located on the top of nasal bridge and should possess very small SN_y value. Therefore, the candidate root regions are relatively flat on the SN_y map, which can be directly detected from the profile



(a) Nasal profile signature on the depth map: the blue curve is the original profile signature, whose clockwise (red) and counter clockwise (green) rotation are generated to localise the root and subnasal, respectively.



(b) Nasal profile signature on the SN_y map. The nasal root is located on the top of nasal bridge, which is relatively flat. The nose tip presents as the inflexion point on this curve. The subnasal locates at the fast decreasing point.



(c) Gradient of the blue curve in (b)

Figure 11: Root and subnasal localisation by facial profile signatures using an example capture from the Bosphorus database.

signature as shown in Figure 11b. Although there is no distinct pattern for the subnasal presented on the SN_y map, the gradient of the SN_y profile signature curve can be used to help localise the subnasal, as illustrated in Figure 11c, where the subnasal is clearly marked as local minimum. Therefore, combined with the position provided by the SN_y map as shown in Figure 11b, the nasal root and subnasal can be localised at points of inflexion in the nasal signature, for example see Figure 11a.

In summary, the steps described above consist of thresholding surface normal maps, detecting the nasal bridge and using the nasal profile signatures. Taken together, they produce a relatively accurate prediction of the position of the nasal root and subnasal.

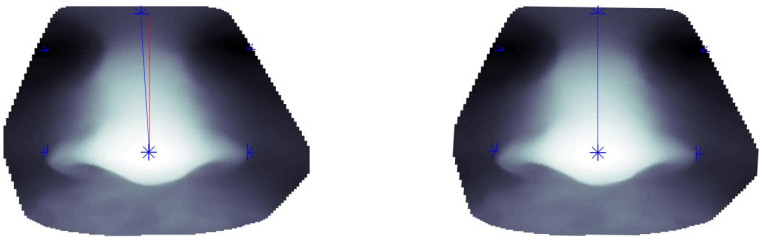
3.3.2. Fine Roll Rotation Adjustment by Using the Detected Landmarks

In Section 3.1.2, the structure of the nasal bridge is used to provide an approximate rotation angle for a rough pose alignment. The detected locations of the nasal root and tip can further be used to correct any remaining pose variations in the roll direction. The example capture presented in Figure 12a has a slight roll rotation that can be effectively detected and corrected by using the locations of the nasal root and tip (shown as the blue points in Figure 12a) to calculate the rotation angle (α_2) between the blue line from root to tip and the y axis (shown in red). As shown in Figure 12b the rotated capture $F_3(x_3, y_3, z_3)$ is found by rotating $F_2(x_2, y_2, z_2)$ by angle α_2 in the x-y plane, where the value of α_2 is given by,

$$\alpha_2 = tan^{-1}\frac{root_x - tip_x}{root_y - tip_y} \tag{12}$$

and the rotated coordinates (x_3, y_3, z_3) are then,

$$\begin{cases} x_3 = x_2 \times cos\alpha_2 - y_2 \times sin\alpha_2 \\ y_3 = x_2 \times sin\alpha_2 + y_2 \times cos\alpha_2 \\ z_3 = z_2 \end{cases} \tag{13}$$



(a) facial capture with slight roll rotation (b) aligned capture

Figure 12: Roll rotation calibration using detected landmarks.

3.4. Eye Corners Detection

3.4.1. Eye Regions Localization by Thresholded SNs

In the first stage of the eye regions localisation is to find their approximate horizontal positions by thresholding the SN maps. The concave regions on both sides of nasal root, shown by the white lines in Figure 13a, possess a geometric structure that is distinct from other regions around the nasal root in both the depth and SN maps. Instead of calculating and thresholding the curvatures, simply thresholding the SN can also produce reliable candidate eye corner (CEC) positions. These regions have higher values of SN_x and lower values of SN_y. It is therefore proposed to find the CEC regions shown in Figure 13b using

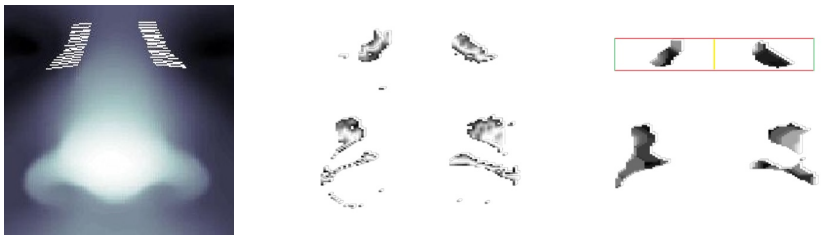
$$CEC = \{C_{i,j} | C_{i,j} = f_{x,y}(i,j)\} \tag{14}$$

where, given SN_x and SN_y,

$$f_{x,y}(i,j) = \begin{cases} 1 & if\ (SN_x(i,j) > T_{ex}) \wedge (SN_y(i,j) < T_{ey}) \\ 0 & otherwise \end{cases} \tag{15}$$

where T_{ex} and T_{ey} and thresholds.

As shown in the example in Figure 13b, because of the noisy points, the detected CEC regions may not be fully connected in the thesholded SN maps. Therefore, the morphological closing operation is first applied to the initial CEC map using a disk-shaped structuring element with radius 5. A morphological



(a) (b) (c)

Figure 13: 75 nasal candidate eye corner localised by the thresholded SNs. The target concave region around the root is shown in (a). After thresholding the surface normals in (b), the candidate eye corners are shown in (c).

erosion is used to remove some small noisy points. Figure 13c illustrates the final candidate region after the morphological post-processing steps.

3.4.2. Corner Localization

Using the thresholded surface normals maps (*CEC* regions) and also the reference points provided by the previously detected root and alar grooves, the final eye corner regions can be localised in the rectangular regions shown in Figure 13c. To exclude most irrelevant regions, the position of the nasal root is used to find the top and bottom boundaries (red lines) of eye corner regions. The left and right boundaries (green lines) are set by the position of two alar grooves in the x axis. Therefore, the eye corner regions shown within the rectangles are successfully cropped from the thresholded surface normals maps. Within the candidate eye corner region of each side, the eye corner is located at the global minimum of the depth map.

4. Landmarking Performance Evaluations

In this section, the proposed landmarking algorithm is evaluated by measuring the within-class consistency and accuracy using the Bosphorus, FRGC and BU-3DFE databases. These databases are selected as they are challenging, containing a large number of captures with different expressions, and are

widely used in 3D landmarking evaluation. The FRGC database has three folders (Spring2003, Fall2003 and Spring2004) and all its 4950 samples are used as they contain the widest variation in expressions. Excluding the captures with ground-truth pose variations and occlusions, captures with various expressions of 105 subjects from the Bosphorus database are used. Some Bosphorus captures with occlusions of the non-nasal region are also used, as are those with pose variations within 10°, which can be adjused by the proposed landmarking algorithm. BU-3DFE is a dedicated expression database, which has one neutral and four levels (low, middle, high and highest) of 6 expressions (anger, happiness, sadness, surprise, disgust and fear). This gives 25 3D expressions for each subject and results in a database of 2,500 3D facial expressions.

In the landmarking algorithm performance evaluations performed below, the same threshold values for the SN maps are used for all three databases. Compared with many previously proposed landmarking algorithms, for example [10] and [26], the proposed approach is computationally efficient, completely training free and can be generalized to any 3D face databases, containing high resolution facial information. Using MATLAB 2017b on a 2.8 GHz Intel Core i7 with 16G RAM, the average landmarking time is 0.025 seconds per capture.

T_x is set to 0.1 for nasal bridge detection in Equation 2 and T_z is set to 0.3 for the candidate alar grooves thresholding (Equation 6). The use of i is to crop the candidate regions of alar grooves and the range of 5-21 was tested. For the consistency and accuracy evaluations, similar performances are obtained for i = 11 to 21 and the robustness reduces when $i < 11$. Therefore, i is set to 17 for the proposed landmarking algorithm. The two thresholds, T_{ex} and T_{ey}, used for detecting the candidate eye corner positions in Equation 15 are set to 0.7 and 0.15, respectively. The same settings can be generalized to most types of 3D face data, without changing the parameters.

4.1. Within-class Consistency Evaluation

Although the Bosphorus, FRGC and BU-3DFE databases provide some manually detected landmarks, finding the ground truth locations for these land-

marks is still very challenging. Evaluating the within-class consistency is a useful strategy, which can be obtained by calculating the standard deviation [2]. In this process, all the captures are pose aligned and translated to the nose tip. For each given landmark, the standard deviations of the x, y and z positions are first estimated and the average standard deviation (in mm) of all the subjects is measured as the consistency. As the majority of captures in the Bosphorus database contain artificial expressions, including 6 basic expressions and different Face Action Units, those captures are used to further evaluate the within-class consistency of different types of expressions.

4.1.1. All Captures Scenario

In this scenario, all the captures in three databases are used for within-class consistency evaluation and a comparison between the proposed landmarking algorithm and the recently proposed landmarking scheme of Emambakhsh et al. [2] is shown in Table 1. The within-class consistencies of all the main nasal landmarks have been significantly improved, except the nasal root in the Bosphorus database and the proposed landmarking algorithm is more consistent than that of [2]. This demonstrates that the proposed landmarking algorithm has the ability to deal with more complicated expression variations.

Many of the captures in the FRGC database contain natural expressions and so it has less variations in expressions than the Bosphorus database. However, there are small pose variations in most captures, which is another big challenge in facial landmarking. Compared to the results shown in [2], the consistencies of the FRGC landmarks are significantly improved, which demonstrates that the proposed landmarking algorithm has the ability to deal with most small pose variations. The consistency of the detected landmarks provided by Cruesot et al. [10] were also calculated but their results were found to be much worse than those of both methods evaluated in Table 1. The within-class consisteny was also evaluated on the BU-3DFE database, giving results of 1.06 $\pm$ 0.35 (Root), 0.59 $\pm$ 0.38 (Alar_L), 0.60 $\pm$ 0.37 (Alar_R), 1.36 $\pm$ 0.27 (Eye_L), 1.39 $\pm$ 0.25 (Eye_R) and 0.58 $\pm$ 0.10 (Subnasal).

Database	Landmarks	Emambakhsh et al.[2]	Proposed
Bosphorus	Root	1.06 ± 0.58	1.27 ± 0.27
	Alar_L	1.06 ± 0.62	0.80 ± 0.21
	Alar_R	1.19 ± 0.60	0.78 ± 0.22
	Eye_L	1.76 ± 1.03	1.51 ± 0.34
	Eye_R	2.12 ± 1.14	1.47 ±0.32
	Subnasal	1.11 ± 0.38	1.09 ± 0.22
FRGC	Root	2.04 ± 1.09	0.81 ± 0.45
	Alar_L	1.29 ± 0.82	0.50 ± 0.31
	Alar_R	1.22 ± 0.62	0.51 ± 0.33
	Eye_L	2.95 ± 1.61	1.15 ± 0.60
	Eye_R	2.91 ± 1.53	1.17 ± 0.61
	Subnasal	1.86 ± 0.85	0.78 ± 0.50

Table 1: A comparison of within-class landmarks consistency in mm using the method proposed in [2]. All the captures from Bosphorus and FRGC databases are pose aligned and translated to the nose tip and the average and standard deviations of six main nasal landmarks for all the subjects are illustrated.

4.1.2. Different Types of Expressions

Using all the captures in the Bosphorus database for the landmarks within-class consistency evaluation can generally show whether the target landmark is localised to the same point or area for each capture of the same identity. In fact, for an ideal landmarking algorithm, the consistency of each landmark under different expressions is expected to vary. Specifically, the consistency between the neutral captures and those captures with very small variations on the nasal region due to expressions should be better than those containing large expression variations. This is because more severe expressions will cause more facial surface movement. Therefore, it is useful to further evaluate the within-class consistency of the root (upper) and alar grooves (lower) using different types of expressions.

To do this, the Bosphorus database is used as it contains a wide range of facial expressions. According to the database labels provided, all the captures from 105 subjects are divided into 34 groups, including 6 basic expressions, up-

Expressions	Root	Alar_L	Alar_R	Expressions	Root	Alar_L	Alar_R
CAU_A12A15_0	1.81	1.75	1.45	LFAU_17_0	1.42	1.60	1.39
CAU_A22A25_0	1.33	1.30	1.05	LFAU_18_0	1.57	1.40	1.28
CAU_A26A12lw_0	1.29	1.18	1.15	LFAU_20_0	1.63	1.40	1.40
E_ANGER_0	2.02	1.27	1.13	LFAU_22_0	1.66	1.33	1.24
E_DISGUST_0	2.68	1.57	1.69	LFAU_23_0	1.64	1.33	1.27
E_FEAR_0	1.55	1.33	1.14	LFAU_24_0	1.50	1.32	1.26
E_HAPPY_0	1.54	2.67	2.29	LFAU_25_0	1.18	1.10	1.07
E_SADNESS_0	1.79	1.63	1.30	LFAU_26_0	1.29	1.31	1.13
E_SURPRISE_0	1.40	1.25	1.27	LFAU_27_0	1.69	1.28	1.35
LFAU_10_0	2.26	1.95	1.84	LFAU_28_0	2.15	1.57	1.45
LFAU_12LW_0	1.28	1.07	1.39	LFAU_34_0	1.76	1.28	1.27
LFAU_12L_0	1.34	1.41	1.57	LFAU_9_0	4.52	1.67	2.03
LFAU_12R_0	1.47	1.44	1.34	UFAU_1_0	2.05	1.18	1.31
LFAU_12_0	1.50	1.95	2.26	UFAU_2_0	1.61	1.15	1.04
LFAU_14_0	1.44	1.53	1.32	UFAU_43_0	1.52	1.12	1.14
LFAU_15_0	1.46	1.63	1.27	UFAU_44_0	1.97	1.16	0.98
LFAU_16_0	1.30	1.11	0.99	UFAU_4_0	2.25	1.31	1.28

Table 2: Within-class landmarks consistency evaluation using different types of expressions. All the captures contain the expression variations from the Bosphorus database are divided into 34 groups and the average distance of the landmarks for all the subjects in each group is calculated. The Euclidean distance is used to measure the distance between two landmarks on the neutral and non-neutral captures.

per and lower Face Action Units (UFAU and LFAU) and some combined Action Units (CAU). The neutral capture of each identity is applied as a reference and each identity has only one capture for each group. Therefore, the consistency of each landmark is measured by the Euclidean distance between the same landmarks on the neutral and non-neutral captures of the same identity. The average distance of the root and alar grooves for each group is presented in Table 2.

Compared with the overall within-class consistency, calculated by the average standard deviations, the results shown in Table 2 indicate the different expressions have varying influence on the consistency of the landmarks within the nasal region. In some the groups, for example LFAU_25_0, the nasal landmarks are more consistent as less facial movement is caused by these expressions. Most UFAU groups, which cause the nasal root area to stretch or extrude, show

a poor consistency for the nasal root but a good consistency for the alar grooves. In contrast, LFAU groups have a poor consistency for the alar grooves but a good consistency for the root, especially for the captures containing the happy expression. Considering all different types of facial expressions confirms that the proposed landmarking algorithm demonstrates a good consistency when the captures contain less severe expressions. For captures with large expression variations within the nasal region, it is better to evaluate the landmark consistency combined with the accuracy.

4.2. Accuracy

Calculating the accuracy of each landmark is another widely used method to evaluate the performance of a landmarking algorithm. To do this, the locations of the landmarks are compared with the ground truth in the Bosphorus and FRGC databases provided by Savran et al. [7], Creusot et al. [10] and the ground truth provided by the BU-3DFE database [9]. As the nasal root, tip and subnasal are not provided in the BU-3DFE database, only the alar grooves and eye corners are considered.

The definitions of the nasal tip, root and subnasal in the proposed landmarking algorithm consider the surface normals together with the depth information, which means some minor surface characteristics can be detected and requires more accurate manual annotations. Although the definitions of those nasal landmarks are not necessarily consistent with the manual annotations provided in [10], the manually annotated landmarks still provide a good reference for landmarking precision evaluation.

However, as shown in Figure 14, the definition of alar grooves is more problematic than for other landmarks. Figure 14a and 14b shows inconsistent manual marking of ground truth positions in BU-3DFE. While many landmarking algorithms try to localise the alar grooves positions marked in Figure 14a [shown as green points in Figure 14c and 14d], in the proposed method they are defined as the outermost points, consistent with Figure 14b and shown as red points in Figure 14c and 14d. The main motivation of using the outermost points is their

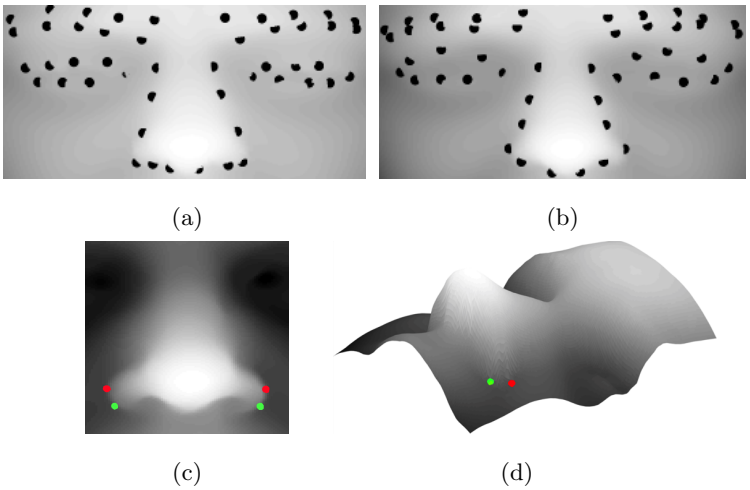


Figure 14: Example ground truth landmarks from the BU-3DFE database and two different alar grooves locations. (a) and (b) Two facial captures with ground truth landmarks from the same identity showing the inconsistent annotation of the nasal alar landmarks. (c) and (d) Two views of the alar grooves showing the landmarks defined in the proposed method (red points) and those used in many published works (green points).

consistency and simplicity. Compared to the other positions used, the outermost points are easy to localized by their local patterns and the reference prediction provided by the nose tip. Furthermore, they can be consistently localized, as shown in Table 1, where the consisencies of two alar grooves are siginificantly better than the other nasal landmarks. Similarly, for BU-3DFE the within-class consistency for the alar grooves was found to be the lowest of all 6 landmarks, apart from the subnasal.

The mean landmarking error for each main nasal landmarks is calculated on the FRGC, Bosphorus and BU-3DFE captures and shown in Table 3, 4 and 5, respectively. Compared with the state-of-the-art algorithms, the proposed low complexity landmarking algorithm has the lowest landmarking errors for the tip, root and subnasal under expression and slight pose variations. The mean errors of two eye inner corners are very competitive, beating the majority of comparison techniques for all databases, and only a little higher than [26].

The mean errors results for the alar grooves are worse than other works as

Method	Tip	Root	Subnasal	Alar_L	Alar_R	Eye_L	Eye_R
Segundo et al.[15]	2.73			4.83	5.84	3.69	3.35
Cruesot et al.[10]	3.36	4.22	3.65	3.72	3.72	5.93	5.93
Sukno et al.[25]	2.22	2.55	2.81	3.09	3.09	4.49	4.49
Gilani et al.[26]	2.2	2.5	3.4	3.0	3.0	2.4	2.4
Proposed	1.21	1.89	1.56	6.78	6.87	3.33	3.31

Table 3: Comparison of the mean of landmarking errors over the FRGC captures.

Method	Tip	Root	Subnasal	Alar_L	Alar_R	Eye_L	Eye_R
Cruesot et al.[10]	4.47	6.33	5.23	4.22	4.07	5.09	4.18
Sukno et al.[25]	2.33	2.22	2.81	3.03	3.01	2.94	2.76
Gilani et al.[26]	2.82	2.32	2.35	2.5	2.99	2.35	2.40
Proposed	1.49	2.11	1.91	6.74	6.78	3.01	2.98

Table 4: Comparison of the mean of landmarking errors over the Bosphorus captures.

Method	Alar_L	Alar_R	Eye_L	Eye_R
Segundo et al.[15]	6.6	6.6	6.3	6.3
Salazar et al. [30]	6.8	6.8	6.8	6.1
Gilani et al.[26]	2.9	2.9	2.8	2.7
Proposed	4.7	4.9	3.4	3.3

Table 5: Comparison of the mean of landmarking errors over the BU-3DFE captures.

different ground truth positions are used, see Figure 14c and 14d. The proposed technique's results for the BU-3DFE database are better because some of the annotated landmarks are a closer match to the outermost positions it uses.

Figure 14 and the accuracy results reported for most algorithms [15][25][26] in Table 3, 4 and 5, show that the alar grooves are much harder to localilize in comparison with the other nasal landmarks, which implies the ground truth positions of the alars marked in green in Figure 14 are very challenging to localize. In comparison, the outmost points localized by the proposed approach and shown in red in Figure 14 provide a more robust and accurate alternative. Although the mean landmarking error of the proposed approach for the alar grooves is larger than those of other works, its standard deviation is very low. According to the within-class consistency evaluation above, it is evident that the alar grooves have been consistently localized to a point slightly above the

provided ground truth point.

5. Conclusions

In this paper, a novel landmarking algorithm is proposed to improve the consistency and accuracy of the main nasal landmarks. This new landmarking strategy was originally inspired by the successful use of SN for feature extraction, where they have been shown to provide an additional and effective representation for 3D surface geometry structure analysis. The thresholded SN are first used to localise the nasal bridge, which provides a coarse roll correction, and to help refine the position of the nasal tip and correct the yaw rotations. Using the refined tip location, the candidate alar grooves and root regions are found by thresholding the SN. The nasal root and subnasal are localised by the facial profile signature and the alar grooves are found by utilizing the facial transect. Finally, seven nasal landmarks (the tip, root, subnasal, alar grooves and eye corners) are robustly localised on the well-aligned 3D nasal region. This new landmarking strategy provides an effective, low complexity way to localise the key landmarks on human nose and is tested on the Bosphorus, FRGC and BU-3DFE databases which contain a large number of expression variations. The results show that the detected landmarks demonstrate good within-class consistency and accuracy.

The use of surface normals for landmarking in 3D faces has therefore been shown to have a number of benefits. The proposed approach is low complexity and its results for several key landmarks outperform those produced by more complex state-of-the-art approaches, without the need for training. Areas of further work that would be interesting to include investigating the robustness of the proposed approach to more severe pose variations and also to low resolution, noisy images. For the latter, as surface normals have similar properties to other 3D representations such as curvatures or shape index, their stability performance may well prove to be equivalent.

References

[1] X. Yu, Y. Gao, J. Zhou, Sparse 3D directional vertices vs continuous 3D curves: Efficient 3D surface matching and its application for single model face recognition, Pattern Recognition 65 (2017) 296–306.

[2] M. Emambakhsh, A. Evans, Nasal patches and curves for an expression-robust 3D face recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 39 (5) (2017) 995–1007.

[3] S. M. Islam, R. Davies, M. Bennamoun, R. A. Owens, A. S. Mian, Multibiometric human recognition using 3D ear and face features, Pattern Recognition 46 (3) (2013) 613–627.

[4] J. Shi, A. Samal, D. Marx, How effective are landmarks and their geometry for face recognition?, Computer Vision and Image Understanding 102 (2) (2006) 117–133.

[5] J. Gao, A. N. Evans, Expression robust 3D face recognition by matching multi-component local shape descriptors on the nasal and adjoining cheek regions, in: 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), Vol. 1, IEEE, 2015, pp. 1–8.

[6] H. Li, D. Huang, J.-M. Morvan, L. Chen, Y. Wang, Expression-robust 3D face recognition via weighted sparse representation of multi-scale and multi-component local normal patterns, Neurocomputing 133 (2014) 179–193.

[7] A. Savran, N. Alyüz, H. Dibeklioğlu, O. Çeliktutan, B. Gökberk, B. Sankur, L. Akarun, Bosphorus database for 3D face analysis, in: European Workshop on Biometrics and Identity Management, Springer, 2008, pp. 47–56.

[8] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, W. Worek, Overview of the face recognition grand challenge, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1, IEEE, 2005, pp. 947–954.

[9] L. Yin, X. Wei, Y. Sun, J. Wang, M. J. Rosato, A 3D facial expression database for facial behavior research, in: 7th International Conference on Automatic Face and Gesture Recognition, 2006, pp. 211–216.

[10] C. Creusot, N. Pears, J. Austin, A machine-learning approach to keypoint detection and landmarking on 3D meshes, International Journal of Computer Vision 102 (1-3) (2013) 146–179.

[11] K. I. Chang, K. W. Bowyer, P. J. Flynn, Multiple nose region matching for 3D face recognition under varying facial expression, IEEE Transactions on Pattern Analysis and Machine Intelligence 28 (10) (2006) 1695–1700.

[12] D. Colbry, G. Stockman, A. Jain, Detection of anchor points for 3D face veri. cation, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops, IEEE, 2005, pp. 118–118.

[13] J. D'Hose, J. Colineau, C. Bichon, B. Dorizzi, Precise localization of landmarks on 3D faces using gabor wavelets, in: First IEEE International Conference on Biometrics: Theory, Applications, and Systems, IEEE, 2007, pp. 1–6.

[14] N. Pears, T. Heseltine, M. Romero, From 3D point clouds to pose-normalised depth maps, International Journal of Computer Vision 89 (2-3) (2010) 152–176.

[15] M. P. Segundo, L. Silva, O. R. P. Bellon, C. C. Queirolo, Automatic face segmentation and facial landmark detection in range images, IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 40 (5) (2010) 1319–1330.

[16] P. Szeptycki, M. Ardabilian, L. Chen, A coarse-to-fine curvature analysis-based rotation invariant 3D face landmarking, in: IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems, IEEE, 2009, pp. 1–6.

[17] A. Colombo, C. Cusano, R. Schettini, 3D face detection using curvature analysis, Pattern Recognition 39 (3) (2006) 444–455.

[18] J. J. Koenderink, A. J. Van Doorn, Surface shape and curvature scales, Image and Vision Computing 10 (8) (1992) 557–564.

[19] X. Lu, A. K. Jain, Automatic feature extraction for multiview 3D face recognition, in: 7th International Conference on Automatic Face and Gesture Recognition, IEEE, 2006, pp. 585–590.

[20] T. C. Faltemier, K. W. Bowyer, P. J. Flynn, Rotated profile signatures for robust 3D feature detection, in: 8th IEEE International Conference on Automatic Face & Gesture Recognition, IEEE, 2008, pp. 1–7.

[21] S. Berretti, A. Del Bimbo, P. Pala, Recognition of 3D faces with missing parts based on profile networks, in: Proceedings of the ACM workshop on 3D object retrieval, ACM, 2010, pp. 81–86.

[22] X. Zhao, E. Dellandrea, L. Chen, I. A. Kakadiaris, Accurate landmarking of three-dimensional facial data in the presence of facial expressions and occlusions using a three-dimensional statistical facial feature model, IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 41 (5) (2011) 1417–1428.

[23] P. Perakis, G. Passalis, T. Theoharis, I. A. Kakadiaris, 3D facial landmark detection under large yaw and expression variations, IEEE Transactions on Pattern Analysis and Machine Intelligence 35 (7) (2013) 1552–1564.

[24] P. Perakis, T. Theoharis, I. A. Kakadiaris, Feature fusion for facial landmark detection, Pattern Recognition 47 (9) (2014) 2783–2793.

[25] F. M. Sukno, J. L. Waddington, P. F. Whelan, 3-D facial landmark localization with asymmetry patterns and shape regression from incomplete local features, IEEE Transactions on Cybernetics 45 (9) (2015) 1717–1730.

[26] S. Z. Gilani, A. Mian, F. Shafait, I. Reid, Dense 3D face correspondence, IEEE Transactions on Pattern Analysis and Machine Intelligence, in press, doi: 10.1109/TPAMI.2017.2725279.

[27] M. Emambakhsh, A. N. Evans, M. Smith, Using nasal curves matching for expression robust 3D nose recognition, in: IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS), IEEE, 2013, pp. 1–8.

[28] K. W. Bowyer, K. Chang, P. Flynn, A survey of approaches and challenges in 3D and multi-modal 3D+ 2D face recognition, Computer Vision and Image Understanding 101 (1) (2006) 1–15.

[29] J. Gao, M. Emambakhsh, A. N. Evans, A low dimensionality expression robust rejector for 3D face recognition, in: 22nd International Conference on Pattern Recognition (ICPR), IEEE, 2014, pp. 506–511.

[30] A. Salazar, S. Wuhrer, C. Shu, F. Prieto, Fully automatic expression-invariant face correspondence, Machine Vision and Applications 25 (4) (2014) 859–879.